AP STAT: CHAPTER 3 **CATEGORICAL DATA**

MAKE A PICTURE!

First, create a frequency table

Example: number of students at CB South in each grade:

Grade	TOTAL
10	534
11	552
12	515

Proportion = decimal 10^{th}: 534/total = 0.30

Percent = % 30%

Frequency = # count

Relative frequency = %

Distribution (of a variable)- values of variable & how often we see each value

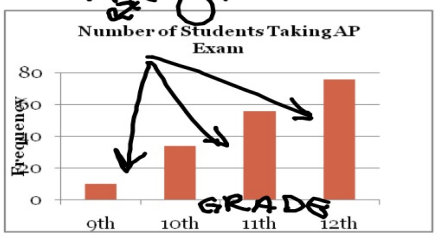
Examples: bar chart, pie chart, stemplot, histogram, boxplot, dotplot

Categorical Distributions:

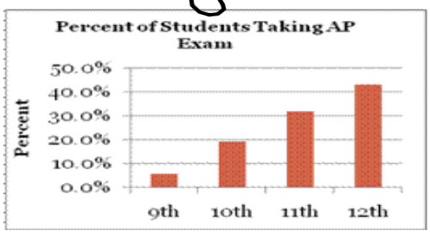
1. Bar Chart

label axes

\# freq.



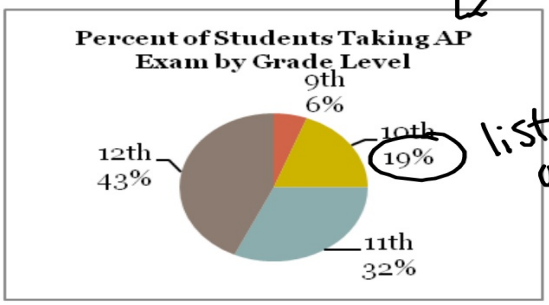
rel. freq.



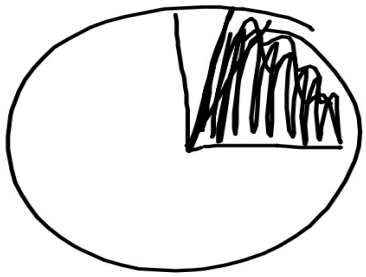
Pie charts

%

label

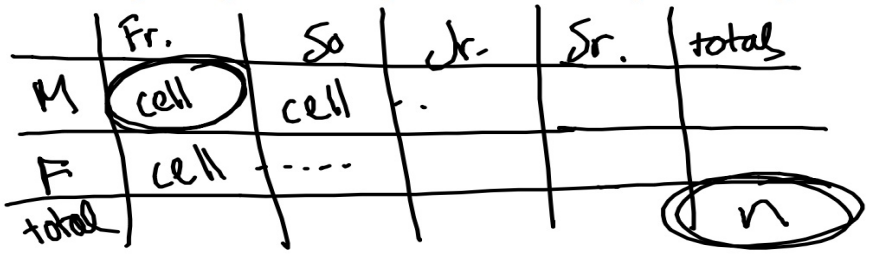


listed as %.



7/20

3. Contingency tables (aka Two-Way tables)



n = 100

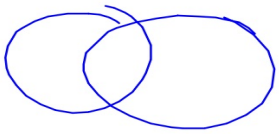
Identify:

- Row variable — Gender
- Column variable — Grade
- Values of the variable — M/F — Fr./So/Jr/Sr
- Total (n)
- # of Cells — 8
- Totals

Example: Hospitals

	Hospital A	Hospital B	
Died	63	16	79
Survived	2037	784	2821
	2100	800	2900

- What percent of people died? $\frac{79}{2900} = 2.72\%$
- Of those people that went to Hospital A, what percent died? $\frac{63}{2100} = 3\%$
- Of those people who went to Hospital B, what percent died? $\frac{16}{800} = 2\%$
- Of those people who died, what percent went to Hospital A? $\frac{63}{79} = 79.75\%$
- What percent of people died and went to Hospital B? 0 overlap $\frac{16}{2900} = 0.55\%$



2 types of Distributions for Categorical Variables — 2 way tables

1) **MARGINAL DISTRIBUTIONS** (totals)

- How to make: margins ÷ n (total)
- Looking for ... overall % in each category value
- ALWAYS ... in %

- Example: Hair color vs. Gender

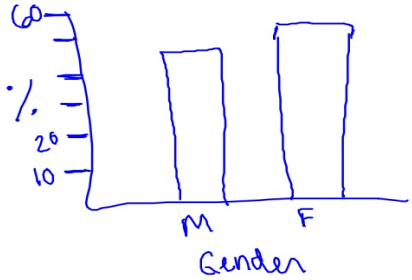
	Brown	Blonde	Black	Red	Total
MALE	26	24	10	3	63
FEMALE	20	35	12	6	73
TOTALs	46	59	22	9	136

- Find the marginal distribution for the HAIR COLOR variable

Brown = 46/136 = 33.82 %
Blonde = 59/136 = 44.38%
Black = = 16.18%
Red = = 6.62%

- Find the marginal distribution for the GENDER variable

Male = 46.32%
Female = 53.68%



- Represented Visually: BAR CHART

1) CONDITIONAL DISTRIBUTIONS

- Look at … one variable

Then look at … each value individually

- Break down … each value into its pieces
- ALWAYS … %

- Example: Hair Color vs. Gender

Hair color

	Brown	Blonde	Black	Red	Total
MALE	26	24	10	3	63
FEMALE	20	35	12	6	73
TOTALs	46	59	22	9	136

- Find the conditional Distribution for the HAIR COLOR variable

Brown
M: 26/46 = 56.52%
F: 20/46 = 43.48%

Blonde
M 24/59
F 35/59

Black
M: 10/22
F: 12/22

Red
M: 3/9 =
F: 6/9 =

- Find the conditional Distribution for the GENDER variable

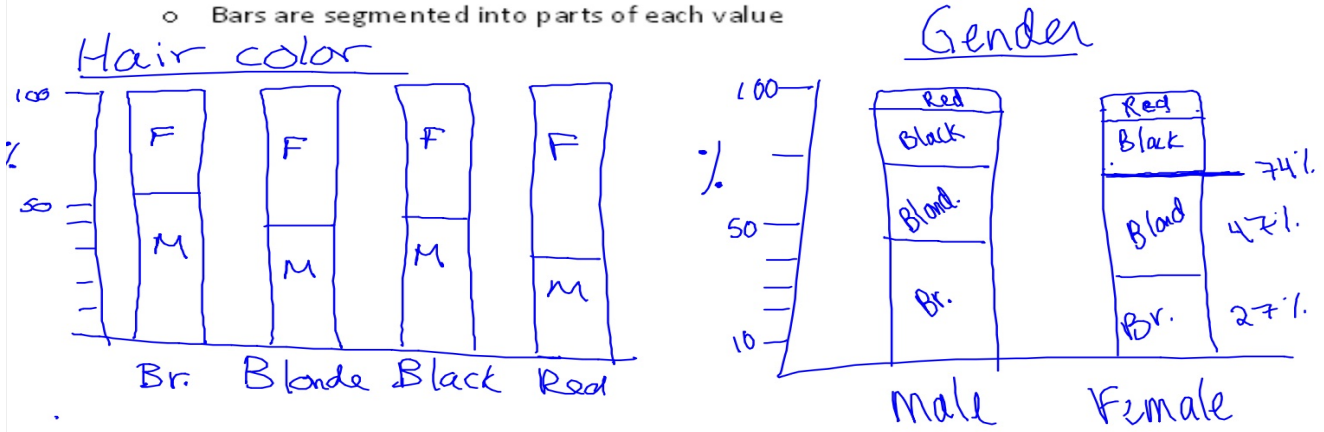
Male

Br.	26/63 =	41.27%
Blon	24/63 =	38.1%
Bla	10/63 =	15.87%
Red	3/63 =	4.76%

Female

Br	20/73 =	27.4%
Blond	35/73 =	47.95%
Bla	12/73 =	16.44%
Red	6/73 =	8.22%

- Represented visually: SEGMENTED (or STACKED) BAR GRAPH
  - Each bar = 100%
  - Values of variable on the x-axis
  - Bars are segmented into parts of each value



Try worksheet 3A on your own! Work with someone else in the class if you want